基于参数优化SVM算法的交联电缆局部放电模式识别

为避免传统模式识别中神经网络收敛速度慢、过学习等不足或因支持向量机（Support Vector Machine，SVM）参数选择不当而导致识别准确度低，在引入M－ary分类理论将泛化及学习能力更强的SVM算法扩展为多类分类器的同时，利用改进遗传算法（Genetic Algorithm，GA）优化各子分类器的惩罚因子、松弛变量及核函数参数，从而构造出最优参数SVM分类模型。开展交联电缆局部放电模拟试验并提取了表征灰度图象内在分形特征的4个分形维数作为局部放电模式识别的放电指纹，分别采用优化SVM、未优化SVM及径向基函数（Radial Basis Function，RBF）神经网络作为模式分类器完成缺陷分类，并通过对比三者识别效果验证了优化模型的正确性。结果表明，以优化SVM作为分类器时各缺陷识别率均大于95%，且无论是否优化参数，SVM总体识别能力要优于RBF神经网络。

胡晓黎　张　皓　段玉兵/山东电力研究院
杨　波　孙晓斌/国网山东省电力公司

胡晓黎/高级工程师

关键词/Keywords
模式识别·
分形特征·
支持向量机·
M－ary分类理论·
参数优化·
遗传算法·

0　引言

近年来，交联聚乙烯（Cross－linked polyethylene，XLPE）电缆因具有合理的工艺结构、安装敷设简单和不存在漏油问题等优势，逐渐成为了我国配电网中电力输送的主流设备，其运行可靠性与电网稳定密切相关。除外界因素，局部放电（以下简称：局放）是导致电缆绝缘故障的主要原因。在对电缆进行局放检测过程中采集到的局放信号承载着电缆全部绝缘故障信息，不同缺陷产生的局放信号特征不同，不同放电类型对电缆绝缘的破坏程度存在巨大差异。因此，电缆局放模式识别问题的研究，对XLPE电缆绝缘诊断有十分重要的意义[1-6]。

截止目前，国内外针对这一问题已开展较多研究。文献[3-4]直接对各种局放$\varphi-q-n$三维谱图及其波形和频谱进行分析比较，虽归纳出了具有一定参考价值的识别规律，但这种识别方法很大程度上取决于工程经验，客观性不强。文献[5]直接将局放脉冲时域波形数据值作为模式识别放电指纹，虽简化了特征提取过程，但即使经过降维处理，识别过程中仍可能遭遇“维数灾难”。文献[6]提取局放灰度图象的分形维数作为神经网络输入，特征量维数适中，取得了不错的识别效果，但人工神经网络方法缺乏数学理论支撑，存在收敛速度慢且易陷入局部极小值的缺点，网络类型及参数选取不当会对分类结果造成很大影响，是一种取决于经验的非线性分类算法。文献[7]利用基于局放统计特征的支持向量机（Support Vector Machine，SVM）完成了变压器故障诊断，但所用的SVM参数单纯依靠人工调试，无法保证识别的速度与准确度。文献[8-9]虽考虑引入遗传算法（Genetic Algorithm，GA）优化SVM参数，但都忽略了松弛变量ξ取值对SVM性能的影响。

基于上述分析，本文采用建立在统计学习理论基础上的新型机器学习算法SVM作为模式识别分类器，在引入M－ary分类理论将SVM扩展为多类分类器的同时，利用改进GA优化各子分类器的惩罚因子C、松弛变量ξ及核函数参数γ，提取局放灰度图象分形维数作为识别特征量，分别利用最优参数组合SVM模型、

未优化SVM分类器及径向基函数（Radial Basis Function，RBF）神经网络识别实验室中模拟的4种XLPE电缆绝缘缺陷。结果表明优化后的SVM分类器具有更高的缺陷识别正确率，且无论是否经过参数优化，SVM算法整体识别率均优于RBF神经网络。

1 利用改进GA实现SVM参数优化

1.1 SVM算法基本原理

对于非线性分类问题，SVM算法基本思想在于：将样本空间中的线性不可分部分非线性映射至高维甚至是无穷维的特征空间中，并在此空间建立一最优分类超平面，使两类训练样本分类间隔达到最大，由此可将分类问题转化为如下凸二次规划问题[10]

$$\begin{cases} \min \dfrac{1}{2}\omega^{\mathrm{T}}\omega + C\sum\limits_{i=1}^{l}\xi_i \\ \text{s.t.} \quad y_i\left[\left\langle \omega \cdot x_i \right\rangle + b\right] \geqslant 1-\xi_i \\ \qquad \xi_i \geqslant 0 \quad i=1,2,\cdots,l \end{cases} \tag{1}$$

式中，(x_i, y_i)为训练样本且$x\in R^{\mathrm{n}}$，$y=\{+1, -1\}$；$\omega\in R^{\mathrm{n}}$为平面法向量；$b\in R$为平面截距；C为惩罚因子，用于控制算法运行复杂程度与逼近误差的折衷，该值过小会因对错分样本惩罚力度不够使算法陷入欠学习，而该值过大则会引起过学习现象[8]；ξ_i为松弛变量，用于容许错分样本存在。

为解决式（1）中的凸二次规划问题，考虑数学中带约束极值问题的求解方法，引入Lagrange乘子$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_l)$，$\alpha_i\geqslant 0$，$(i=1,2,\cdots,l)$，将式（1）转化为其对偶问题

$$\begin{cases} \max \sum\limits_{i=1}^{l}\alpha_i - \dfrac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) \\ \text{s.t.} \quad 0\leqslant\alpha\leqslant C \quad i=1,2,\cdots,l \\ \qquad \sum\limits_{i=1}^{l}\alpha_i y_i = 0 \end{cases} \tag{2}$$

求解上式得最优解α_i^*，使α_i^*不为0的样本就是支持向量。依据泛函理论，在最优分类超平面中引入符合Mercer条件的核函数$K(x_i, x_j)$即可实现非线性映射后的线性分类。故分类决策函数为

$$F(x)=\mathrm{sgn}\left(\sum_{i=1}^{l} y_i\alpha_i^* K(x,x_i)+b^*\right) \tag{3}$$

在本文中，核函数选用Gauss函数，即

$$K(x,x_i)=\exp(-\gamma\|x-x_i\|^2) \tag{4}$$

式中，γ为核函数参数，特征空间VC维与其值选取密切相关。

1.2 GA基本原理

GA是一种建立在生物进化论与遗传变异论基础上，通过模拟自然界生物进化机制实现全局寻优的人工智能优化算法，该算法因具有结构简单、鲁棒性强等优势得以广泛应用[11]。GA主要包含以下步骤[8]：①编码，将待优化变量以字符串形式表示；②初始化种群，全局寻优以一个包含潜在最优解的种群为基本单位展开，按照“适者生存，优胜劣汰”准则逐代产生更优个体；③适应度函数，通过计算各个体对应的适应度值挑选出优胜个体；④交叉操作，交换两个体的基因信息产生新个体；⑤变异操作，以小概率随机改变个体基因位点信息。

1.3 SVM参数优化算法设计

考虑到若SVM参数选择不当会对模式识别准确度及算法的运行速度产生负面影响，故本文将GA与SVM算法相结合，利用GA对惩罚因子C、松弛变量ξ及核函数参数γ进行优化，意在构造出最优参数组合SVM模式分类器用于XLPE电缆局放模式识别，SVM参数优化算法具体运行步骤如下[8-9,11-15]。

1）将样本数据分为训练样本与待测样本两部分并进行数据归一化处理。采用浮点编码方式编码待优化参数组合（C，ξ，γ），确定各参数寻优区间。

2）初始化种群设定GA参数，即种群规模N、最大迭代次数T、交叉概率P_{c}及变异概率P_{m}。为避免GA在寻优过程中陷入局部极值，本文引入自适应P_{c}及P_{m}[8]

$$P_c(t+1)=P_c(t)-\frac{[P_c(1)-0.6]}{T} \quad (5)$$

$$P_m(t+1)=P_m(t)-\frac{[0.08-P_m(1)]}{T} \quad (6)$$

式中，t为进化次数，初始值$t=1$。

3）将训练样本识别正确率作为GA寻优的目标函数，计算每一个体对应的函数值并记录最大值。

4）判断是否满足算法终止条件，即迭代次数达到T或目标函数值满足准确度要求。若达到则退出算法并输出参数组合；若未达到则令$t=t+1$对训练个体进行选择、交叉及变异操作后返回步骤3）直到满足条件为止，输出解码即为最优参数组合。

2 XLPE电缆局放信号分形特征提取

2.1 电缆缺陷模型设计及模拟局放试验

XLPE电缆局放形式与缺陷类型密切相关，为客观反映电缆在制作、敷设及运行期间暴露出的典型缺陷的特征与放电机理，本文在4个电缆头中分别制作4类典型缺陷模型：线芯表面毛刺、绝缘内含气泡、局部损伤和绝缘受潮[6]。图1为电缆绝缘缺陷示意图，实际制作中4种缺陷相互独立，互不干扰，图2为其中两种模型实物。

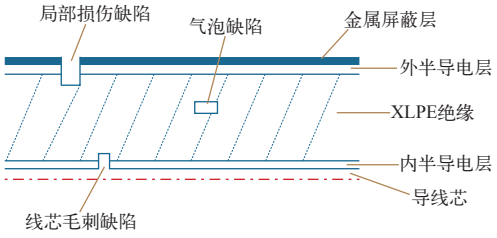
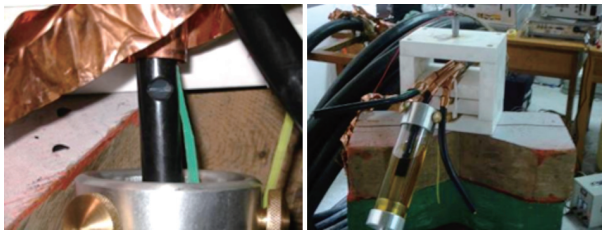


图1 XLPE电缆典型绝缘缺陷示意图



（a）局部损伤模型　　（b）内含气泡模型

图2 局部损伤及内含气泡缺陷模型实物

参照IEC 60270标准，模拟试验选用脉冲电流法采集局放信号，图3为XLPE电缆局放模拟试验接线示意图。为屏蔽通信电磁干扰，获取真实局放试验数据，在全电磁屏蔽实验室中开展试验，图4为局放模拟试验现场。采集局放样本数据时，对不同的缺陷模型采用逐步升压试验法，直到出现稳定的局放现象。每条样本为连续50个工频放电周期累加，每种缺陷模型各测100组样本，共计400组。

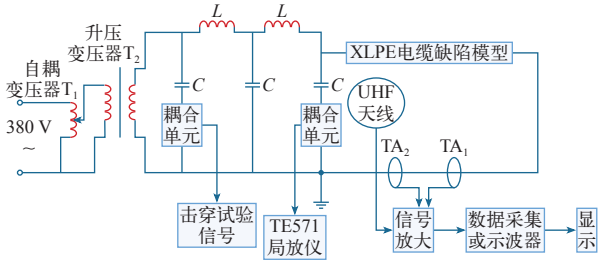


图3 XLPE电缆局放模拟试验接线示意图

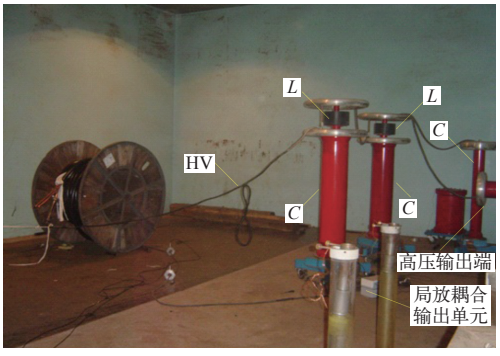


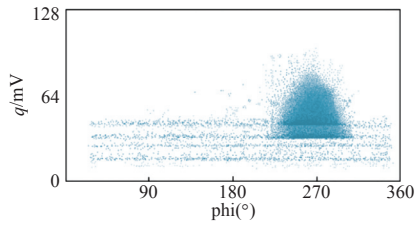
图4 XLPE电缆局放模拟试验现场

2.2 分形维数计算

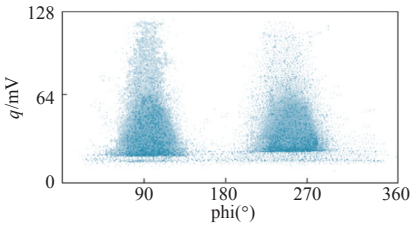
利用计算机处理系统绘制局放$\varphi-q-n$三维谱图，知局放灰度图象灰度级为0～255，依据$\varphi-q-n$谱图中最小放电量与最大放电量分别对应着灰度图象上最小灰度级与最大灰度级的标准，易知各象素点的灰度值为[6]

$$m_{i,j}=\frac{n_{i,j}}{n_{\max}}\times 255 \quad (7)$$

式中，$m_{i,j}$为局放灰度图象上各象素点灰度值；$n_{i,j}$为局放$\varphi-q-n$三维谱图中各小区间内的放电次数；$n_{\max}$为$\varphi-q-n$三维谱图内最大放电次数。图5为对应于线芯表面毛刺缺陷与内含气泡缺陷的局



（a）线芯表面毛刺缺陷时



（b）内含气泡缺陷时

图5　线芯毛刺及内含气泡缺陷时的灰度图象

放灰度图象。

XLPE电缆局放灰度图分形特征盒维数D_B及信息维数D_I的计算分别在图象正、负半周进行，灰度图上灰度值为0的点意味着绝缘状况良好，还没有出现放电现象，故计算中可忽略不计[6]。灰度图分形特征提取流程分别如图6和图7所示。

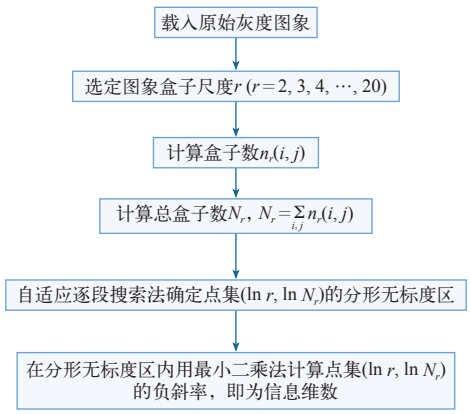


图6　盒维数计算流程

3　XLPE电缆局放模式识别

3.1　SVM分类器扩展

由于将分形维数作为模式分类特征量，故SVM输入样本为四维向量。为将SVM分类器扩展为多类分类器，本文引入M－ary分类思想，首先设定缺陷类型：线芯毛刺（A）、内含气泡（B）、局部损伤（C）及绝缘受潮（D），随后依据M－ary分类原理设计两个二分类SVM子分类器[7]，将对应于4种缺陷类型的输入样本进行重新组合，组合结果如表1所示。

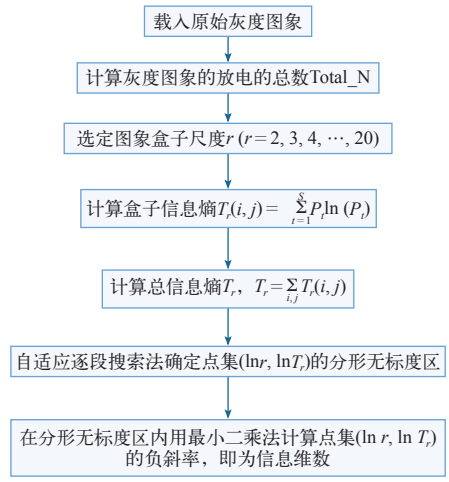


图7　信息维数计算流程

表1　4类缺陷样本数据类型标记

样本标记	SVM_1	SVM_2
+1	A、D	B、D
−1	C、B	C、A

分析上表数据易知，SVM_1将A、D数据组成的样本集标记为正训练样本，SVM_2将C、A数据组成的样本集标记为负训练样本，对于某一条样本数据，若SVM_1识别为正而SVM_2识别为负，则该样本对应的缺陷类型为A类。依此类推，归纳各待识别样本分类归属如表2所示。

表2　待识别样本分类归属

SVM_1	SVM_2	识别结果
+1	+1	D
+1	−1	A
−1	+1	B
−1	−1	C

3.2　RBF神经网络分类器设计

RBF神经网络包含1个输入层、1个输出层和1个带径向基函数的隐含层，具有3层前馈式结构[12]。在本文中，设计RBF网络输入层神经元为4个，与4个分形特征量相对应；输出层神经元为4个，分别对应4种缺陷类型，网络理想输出格式采用二进制来描述；隐含层径向基函数采用高斯函数，设置其

均方误差（MSE）为10^{-6}，该网络能根据输入矢量和输出目标矢量要求自动调整隐含层节点个数，直到满足误差条件为止。

3.3 模式识别结果及分析

在模式识别过程中，4种缺陷类型各有100条有效样本，其中50条用于分类器训练，50条用于测试。首先用训练样本集训练RBF神经网络，按照3.1节方法标记200条训练样本，在进行SVM分类器扩展的同时，利用改进GA分别优化SVM_1及SVM_2参数。优化过程中，设定$N=100$，$T=100$，$P_c(1)=0.8$，$P_m(1)=0.008$，寻优区间$C_{bound}=[0.01, 100]$，$\xi_{bound}=[0, 1]$，$\gamma_{bound}=[0.01, 1\ 000]$。最终得到SVM分类器最优参数组合（$C$，$\xi$，$\gamma$）为$SVM_1$：(81.4，0.01，0.625)，$SVM_2$：(83.8，0.02，0.487)。将待测样本输入分类器，识别结果对比如表3所示。

表3 XLPE电缆局放模式识别结果对比

分类器	正确识别率(%)				整体识别率(%)
	A	B	C	D	
RBFNN	87.8	86.7	90.0	91.3	88.9
SVM	91.2	90.2	92.5	93.3	91.8
GA－SVM	95.3	95.2	97.0	98.3	96.5

分析数据易知，以RBF神经网络作为分类器时，虽局部损伤及绝缘受潮缺陷识别率大于90%，但整体识别率为三者中最低。改用SVM算法分类后，各类缺陷识别率均高于90%且提高至少2.2%，总体识别率提高3.26%。引入改进GA进行参数优化后，GA－SVM分类器对各缺陷识别率均达到了95%以上，与传统SVM相比，各类缺陷识别率提高至少4.4%，整体识别率提高5.12%。因此，优化SVM模型具有最佳的识别正确率，且无论是否优化，使用SVM作为分类器时取得的识别结果整体优于RBF神经网络。

4 结束语

1）本文针对XLPE电缆模式识别中出现的多缺陷分类问题，利用表征局放内在分形特征的分形维数、改进GA及SVM算法，提出了一种基于改进GA优化SVM参数实现XLPE电缆缺陷诊断的新方法。

2）利用所得的最优参数组合SVM分类模型对实验室模拟的4种缺陷类型进行识别。识别结果表明，各类缺陷识别正确率均高于95%，初步验证了该模型的可行性及有效性。

3）将该结果与RBF神经网络及未优化SVM作为分类器时所得诊断结果进行比较。结果表明，与采用RBF神经网络时相比，无论是否进行参数优化，使用SVM作为分类器时取得的识别结果整体优于RBF神经网络；引入GA－SVM分类模型后，与其他两种情况相比，缺陷识别正确率得到了进一步提高，验证了参数优化SVM分类器具有更高的分类准确度。

参考文献

[1] 姚林朋，王辉，钱勇，等．基于半监督学习的XLPE电缆局部放电模式识别研究 [J]．电力系统保护与控制，2011，39 (14)：40-46．

[2] 姚林朋，徐颖，钱勇，等．基于关联规则的XLPE电缆局部放电模糊识别研究 [J]．电工技术学报，2012，27 (5)：92-98．

[3] 朱国正，黄超，毛俊，等．10 kV电缆脉冲电流法局部放电检测技术研究 [J]．华东电力，2014，42 (9)：1876-1881．

[4] 蒋晓娟，何伟，谭延佳，等．110 kV及以上XLPE电缆局放信号的识别方法 [J]．华东电力，2013，41 (6)：1287-1291．

[5] 郑重，谈克雄，王猛，等．基于脉冲波形时域特征的局部放电识别 [J]．电工电能新技术，2001，20 (2)：20-24．

[6] 廖瑞金，犹登亮，周泉．交联聚乙烯电缆局部放电灰度图像的模式识别 [J]．高压电器，2007，43 (2)：85-88．

[7] 褚鑫，张建文，韩刚．统计特征参数及多分类SVM局部放电类型识别 [J]．电测与仪表，2015，52 (7)：35-39．

[8] 刘东平，单甘霖，张岐龙，等．基于改进遗传算

(下转第87页)

例识别出被测装置功能缺陷的过程，说明了本文方法设计的测试用例是合理的。

计量终端自动化功能测试系统到货验收功能测试过程中尚存在检测项的设计必要性和技术依据，测试方法的标准化和规范化等问题有待进一步研究。

参考文献

[1] 王博．实验室计量终端综合测试平台的开发与应用 [J]．自动化应用，2014 (12)：92-93．

[2] 陈佳鹏．基于业务逻辑的移动终端软件自动化测试方法研究与应用 [D]．广州：中山大学，2009．

[3] 邓迎秋．Mac OS X平台移植软件测试技术的研究与实现 [D]．西安：西北工业大学，2007．

[4] 李岩．C++程序的单元测试系统的研究与实现 [D]．沈阳：沈阳工业大学，2007．

[5] 杨建光，梅大成．黑盒测试方法和综合策略的研究 [J]．计算机光盘软件与应用，2012 (4)：121-122．

[6] 帅丹文．接口测试用例设计思路 [J]．程序员，2010 (5)：90-91．

[7] 胡良平，郭辰仪．用SAS软件实现析因设计定量资料的统计分析 [J]．药学服务与研究，2012，12 (4)：249-251．

[8] 周飞，杨根兴，蔡立志．基于UML的测试用例生成方法研究 [J]．计算机应用与软件，2009，26 (2)：107-110．

[9] 中国南方电网有限责任公司．Q/CSG 11109007—2013 中国南方电网有限责任公司配变监测计量终端技术规范 [S]．北京：中国电力出版社，2013．

[10] 畲凤．浅析黑盒测试之决策表法测试技术 [J]．福建电脑，2014 (10)：85-86．

[11] 中国南方电网有限责任公司．Q/CSG 11109004—2013 中国南方电网有限责任公司计量自动化终端上行通信规约 [S]．北京：中国电力出版社，2013．

（收稿日期：2016－02－27）EA

(上接第82页)

法的支持向量机参数优化 [J]．微计算机应用，2010，31 (5)：11-15．

[9] 尹金良，朱永利．支持向量机参数优化及其在变压器故障诊断中的应用 [J]．电测与仪表，2012，49 (5)：11-15．

[10] 李新燕．支持向量机在电缆故障分类中的应用 [D]．西安：西安科技大学，2012．

[11] 王琼瑶，何友全，彭小玲．基于改进遗传算法的支持向量机参数优化方法 [J]．计算机与现代化，2015 (3)：33-36．

[12] 许新征．一种结构自适应的径向基函数神经网络 [J]．计算机工程与应用，2007，43 (14)：75-76．

[13] 贾立静，张建文，王传林，等．基于DGA的差分进化支持向量机电力变压器故障诊断 [J]．高压电器，2015，51 (4)：13-18．

[14] 陆梓端，高茂林．基于改进遗传算法的支持向量机参数优化 [J]．研究与开发，2014，12 (3)：25-29．

[15] 宋淑彩，庞慧，丁学均．GA-SVM算法在文本分类中的研究 [J]．计算机仿真，2011，28 (1)：222-225．

（收稿日期：2015－09－10）EA